Quantifying Digital Mobility Outcomes: A Comprehensive and Objective Methodology for Algorithms' Comparison and Ranking across Datasets

Encarna Micó-Amigo[1], Anne-Elie Carsin[2], Sarah Koch[2], Tecla Bonci[3], Andrea Cereatti[4], Rana Zia Ur Rehman[1], Cameron Kirk[1], Lynn Rochester[1,5], Aida Aydemir[6], Claudia Mazzà[3], Judith Garcia-Aymerich[2], Silvia Del Din[1]

1.Translational and Clinical Research Institute, Newcastle University, UK. 2 Barcelona Institute for Global Health (ISB Global), Spain. 3. Department of Mechanical Engineering & INSIGNEO, The University of Sheffield, UK. 4. Department of Electronics and Telecommunications, Politecnico di Torino, Italy. 5 The Newcastle upon Tyne NHS Foundation Trust, UK. 6. EMD Serono, Billerica, MA, United States, a business of Merck KGaA, Germany

BAM_Research **Maria.Mico-Amigo@Newcastle.ac.uk**

Introduction

- Quantifying **digital mobility outcomes** (DMOs, e.g. walking speed) with wearable sensors (WS) requires the implementation of algorithms from which DMOs are extracted.
- To determine **algorithm validity**, DMOs are compared to equivalent outcomes derived from a gold standard system. However standard methods do not exist.[1]

Aim

1. To develop a comprehensive statistical framework for evaluation of DMOs criterion validity.
2. To present an application of the framework for evaluating performance metrics of step duration and initial contact events detection

Methods & Results

- We used available datasets within the scientific Mobilise-D consortium,[2] to develop and implement a **framework to validate a heterogeneous set of spatiotemporal DMOs** (Table 1). Criterion validity of DMOs was determined by comparison to a corresponding reference system, implemented in laboratory or real-world contexts, and in several cohorts.
- Here we present an application of the framework for evaluating criterion validity of step duration and initial contact events, measured with a WS (APDM, Inc., Portland, OR, USA, 128Hz) at the lower back and using Gaitrite walkway system as a reference (ICICLE-GAIT dataset) in two cohorts: Healthy controls and Parkinson's Disease (Table 2).

Cohorts

- Healthy Controls
- Persons with Parkinson's Disease
- Persons with Multiple Sclerosis
- Stroke survivors, Persons with choreic movement disorders

Reference systems

- Walkway Gaitrite
- Combination of sensors on the shank
- Stereophotogrammetric system
- Combination of inertial and distance sensors

Table 2. *Example of the application of the framework to step duration and initial contact events*

	Initial Contact Events										Step Duration			Step Duration								
	Performance measures (0 to 1)										ICC (-1 to 1)			Relative Error Mean (%)					Relative Error SD (%)				
74 persons with Parkinson's Disease (67.8 +- 9.4 years) 70 Healthy Controls (70.1 +- 7.2 years)	Positive Predictive Value (p - Distribution)	Positive Predictive Value (Mean)	Positive Predictive Value (SD)	Positive Predictive Value (Median)	Positive Predictive Value (IQR)	Sensitivity (p - Distribution)	Sensitivity (Mean)	Sensitivity (SD)	Sensitivity (Median)	Sensitivity (IQR)	Interclass Correlation Coefficient (r)	Interclass Correlation Coefficient (lower bound)	Interclass Correlation Coefficient (upper bound)	Mean Step Duration (p - Distribution)	Mean Step Duration (Mean)	Mean Step Duration (SD)	Mean Step Duration (Median)	Mean Step Duration (IQR)	SD Step Duration (p - Distribution)	SD Step Duration (Mean)	SD Step Duration (SD)	SD Step Duration (Median)	SD Step Duration (IQR)
Marked cells criteria	> 0.05	< 0.7		< 0.7		> 0.05	< 0.7		< 0.7		< 0.7	< 0.7	< 0.7	> 0.05	> 30%	> 30%	> 30%	> 30%	> 0.05	> 30%	> 30%	> 30%	> 30%
PD - Algorithm 1	0.001	0.99	0.04	1.00	0.00	0.003	0.89	0.07	0.90	0.08	0.98	0.97	0.99	0.001	0.96	2.06	0.44	0.41	0.001	2.42	4.97	0.90	1.41
PD - Algorithm 2	0.001	0.99	0.03	1.00	0.00	0.001	0.88	0.05	0.89	0.05	0.92	0.88	0.95	0.001	1.02	4.06	0.30	0.31	0.001	2.48	7.40	0.61	1.26
PD - Algorithm 3	0.001	1.00	0.01	1.00	0.00	0.001	0.89	0.07	0.90	0.06	0.88	0.81	0.93	0.001	0.83	3.34	0.19	0.30	0.001	2.24	6.57	0.79	1.46
HC - Algorithm 1	0.001	0.99	0.03	1.00	0.00	0.001	0.92	0.07	0.95	0.11	0.96	0.94	0.97	0.001	1.84	2.51	0.54	2.39	0.001	9.65	14.62	0.91	16.65
HC - Algorithm 2	0.001	0.99	0.05	1.00	0.00	0.001	0.90	0.07	0.90	0.10	0.89	0.84	0.92	0.001	2.38	3.61	0.58	3.00	0.001	11.01	16.66	0.80	26.37
HC - Algorithm 3	0.001	1.00	0.02	1.00	0.00	0.001	0.91	0.08	0.90	0.11	0.79	0.71	0.86	0.001	1.70	4.37	0.20	0.46	0.001	6.77	14.74	0.57	1.39

Table 1. *Proposed framework: list of DMOs and corresponding analyses*

Block	DMO	PERFORMANCE METRICS					CRITERION VALIDITY							PLOTS		
		Sensitivity	Positive Predictive Value	Accuracy	Specificity	F1-score	Absolute Errors (events)	Relative Errors (events)	Absolute Errors (periods, lengths)	Relative Errors (periods, lengths)	Concurrent validity (ICC)	Limit of Agreement (lower and upper bound)	Significant difference (p)	Bland Altman Plots	Scatter Plots for Correlation (r-value)	Histogram Plots
	Expected acceptable values	>0.7	>0.7	>0.7	>0.7	>0.7		< 30%		<30%	>0.7		>0.05			
GAIT SEQUENCE DETECTION (GSD)	Number of Gait Sequences															
	Start of Gait Sequence detected															
	End of Gait Sequence detected															
	Duration of Gait Sequence detected															
WALKING BOUT ASSEMBLY (WA)	Number of Walking Bouts															
	Start of Walking Bouts filtered															
	End of Walking Bouts filtered															
	Duration of Walking Bouts filtered															
STEP DETECTION (SD)*	Initial Contact Events (Time)															
	Step Duration															
	Stride Duration															
LEFT/RIGHT IDENTIFICATION (L-R)*	Laterality															
	Sequence Order															
CADENCE ESTIMATION (CAD)	Cadence (steps/min)															
STRIDE LENGTH (SL)	Mean Stride Length															
REAL WALKING SPEED (RWS)	Walking Speed															
TURNING DETECTION (TD)*	Number of Turns															
	Start of Turns															
	End of Turns															
	Duration of Turns															
	Maximal Angle of Turns															
HEIGHT ESTIMATION (HE)	Elevation Change															
SECONDARY OUTCOMES (SO)*	Final Contact Events (Time)															
	Stance Phase Duration															
	Swing Phase Duration															

Discussion

We provided a comprehensive statistical framework to evaluate the criterion validity of DMOs, which could be adopted as a tool to standardise DMO validation and facilitate analyses of DMOs in future studies. This framework should be considered by professionals, researchers, and engineers and can be implemented in other fields where evaluation of the validity of outcomes is necessary.

References

1. Bonci T. et al. Sensors. 2020;20(22):6509
2. https://www.mobilise-d.eu/

This study was supported by the European Union's Horizon 2020 research and innovation programme and EFPIA via the Innovative Medicine Initiative 2 (**Mobilise-D project**, grant number **IMI22017-13-7-820820**). The authors would like to thank all the participants and assessors of the **ICICLE study**. ICICLE-Gait study was supported by **Parkinson's UK (J-0802, G-1301)** and by the **National Institute for Health Research (NIHR) Newcastle Biomedical Research Center (BRC)** based at Newcastle Upon Tyne Hospital NHS Foundation Trust and Newcastle University **(REC number: 09/H0906/82)**. The work was also supported by the **NIHR/Wellcome Trust Clinical Research Facility (CRF)** infrastructure at Newcastle upon Tyne Hospitals NHS Foundation Trust. All opinions are those of the authors and not the funders.